# The prediction of pouch of Douglas obliteration using offline analysis of the transvaginal ultrasound 'sliding sign' technique: inter- and intra-observer reproducibility

Reid, S.; Lu, Chuan; Casikar, I.; Mein, B.; Magotti, R.; Ludlow, J.; Benzie, R.; Condous, George

*Published in:*
Acta Obstetricia et Gynecologica Scandinavica

*DOI:*
10.1093/humrep/det044

*Publication date:*
2015

*Citation for published version (APA):*
Reid, S., Lu, C., Casikar, I., Mein, B., Magotti, R., Ludlow, J., Benzie, R., & Condous, G. (2015). The prediction of pouch of Douglas obliteration using offline analysis of the transvaginal ultrasound 'sliding sign' technique: inter- and intra-observer reproducibility. *Acta Obstetricia et Gynecologica Scandinavica*, *94*(12), 1237-1246. https://doi.org/10.1093/humrep/det044

**Human Reproduction, Vol.28, No.5 pp. 1237–1246, 2013**
Advanced Access publication on March 12, 2013 doi:10.1093/humrep/det044

human reproduction

ORIGINAL ARTICLE *Gynaecology*

# The prediction of pouch of Douglas obliteration using offline analysis of the transvaginal ultrasound 'sliding sign' technique: inter- and intra-observer reproducibility

**S. Reid[1,*], C. Lu[2], I. Casikar[1], B. Mein[3], R. Magotti[3], J. Ludlow[4], R. Benzie[3], and G. Condous[1]**

[1]Acute *Gynaecology*, Early Pregnancy and Advanced Endosurgery Unit, Nepean Hospital, University of Sydney, Penrith, NSW, Australia [2]Department of Computer Sciences, University of Wales, Aberystwyth, UK [3]Christopher Kohlenberg Department of Perinatal Ultrasound, Nepean Hospital, University of Sydney, Penrith, NSW, Australia [4]Obstetrics and Gynaecology, Royal Prince Alfred Hospital, Sydney, NSW, Australia

*Correspondence address. E-mail: sereid@gmail.com

*Submitted on September 27, 2012; resubmitted on January 22, 2013; accepted on February 4, 2013*

**STUDY QUESTION:** What is the inter-/intra-observer agreement and diagnostic accuracy among gynaecological and non-gynaecological ultrasound specialists in the prediction of pouch of Douglas (POD) obliteration (secondary to endometriosis) at offline analysis of two-dimensional videos using the dynamic real-time transvaginal ultrasound (TVS) 'sliding sign' technique?

**SUMMARY ANSWER:** The inter-/intra-observer agreement and diagnostic accuracy for the interpretation of the TVS 'sliding sign' in the prediction of POD obliteration was found to be very acceptable, ranging from substantial to almost perfect agreement for the observers who specialized in gynaecological ultrasound.

**WHAT IS KNOWN ALREADY:** Women with POD obliteration at laparoscopy are at an increased risk of bowel endometriosis; therefore, the pre-operative diagnosis of POD obliteration is important in the surgical planning for these women. Previous studies have used TVS to predict POD obliteration prior to laparoscopy, with a sensitivity of 72–83% and specificity of 97–100%. However, there have not been any reproducibility studies performed to validate the use of TVS in the prediction of POD obliteration pre-operatively.

**STUDY DESIGN, SIZE, DURATION:** This was a reproducibility study which involved the offline viewing of pre-recorded video sets of 30 women presenting with chronic pelvic pain, in order to determine POD obliteration using the TVS 'sliding sign' technique. The videos were selected on real-time representative quality/quantity; they were not obtained from sequential patients. There were a total of six observers, including four gynaecological ultrasound specialists and two fetal medicine specialists. The study was conducted over a period of 1 month (March 2012–April 2012).

**PARTICIPANTS/MATERIALS, SETTING, METHODS:** The four gynaecological ultrasound observers performed daily gynaecological scanning, while the other two observers were primarily fetal medicine sonologists. Each sonologist viewed the TVS 'sliding sign' video in two anatomical locations (retro-cervix and posterior uterine fundus), i.e. 60 videos in total. The POD was deemed not obliterated, if 'sliding sign' was positive in both anatomical locations (i.e. anterior rectum/rectosigmoid glided smoothly across the retro-cervix/posterior fundus, respectively). If the 'sliding sign' was negative (i.e. anterior rectum/rectosigmoid did not glide smoothly over retro-cervix/posterior fundal region, respectively), the POD was deemed obliterated. Diagnostic accuracy and inter-observer agreement among the six sonologists was evaluated. The same sonologist was also asked to reanalyse the same videos, albeit in a different order, at least 7 days later to assess for intra-observer agreement. A separate analysis of the inter- and intra-observer correlation was also performed to determine the agreement among the four observers who specialized in gynaecological ultrasound. Cohen's $\kappa$ coefficient $<0$ meant that there was poor agreement, 0.01–0.20 slight agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement and 0.81–0.99 almost perfect agreement.

**MAIN RESULTS AND THE ROLE OF CHANCE:** Agreement (Cohen's $\kappa$) between all six observers for the interpretation of the 'sliding sign' for both sets of videos in both regions (retro-cervix and fundus) ranged from 0.354 to 0.927 (fair agreement to almost perfect agreement) compared with 0.630–0.927 (substantial agreement to almost perfect agreement) when only the gynaecological sonologists were included. The overall multiple rater agreement for the interpretation of the 'sliding sign' for both video sets and both regions was Fleiss' $\kappa$ 0.454 ($P$-value $<0.01$) for all six observers and 0.646 ($P$-value $<0.01$) for the four gynaecological ultrasound specialists. The multiple rater agreement for all six or all four observers was higher for the retro-cervical region versus the fundal region (Fleiss' $\kappa$ 0.542 versus 0.370 and 0.732 versus 0.560, respectively). The intra-observer agreement among the six observers for the interpretation of the 'sliding sign' and prediction of POD obliteration ranged from Cohen's $\kappa$ 0.60–0.95 and 0.46–1.0 ($P$-value $<0.01$), respectively. After excluding the fetal medicine specialists, the intra-observer agreement for the interpretation of the 'sliding sign' and the prediction of POD obliteration ranged from Cohen's $\kappa$ 0.71–0.95 and 0.67–1.0, respectively, indicating substantial to almost perfect agreement. When comparing the four gynaecological observers for the prediction of POD obliteration using the TVS 'sliding sign' (after excluding cases with the POD outcome classified as 'unsure' by the observers), the results for accuracy, sensitivity, specificity, positive and negative predictive value were 93.1–100, 92.9–100, 90.9–100, 77.8–100 and 97.7–100%, respectively.

**LIMITATIONS, REASONS FOR CAUTION:** The 'gold standard' for the diagnosis of POD obliteration is laparoscopy; however, laparoscopic data were available only for 24 out of 30 (80%) TVS 'sliding sign' cases included in this study. Although this should not affect the inter- and intra-observer agreement findings, the ability to draw conclusions regarding the diagnostic accuracy of the TVS 'sliding sign' in the prediction of POD obliteration is somewhat limited. In addition, the diagnostic accuracy findings should be interpreted with the caveat that the cases classified as 'unsure' for the prediction of POD obliteration were excluded from the analysis.

**WIDER IMPLICATIONS OF THE FINDINGS:** We have validated the dynamic real-time TVS 'sliding sign' technique for the prediction of POD obliteration, and this simple ultrasound-based test appears to have very acceptable inter-/intra-observer agreement for those who are experienced in gynaecological ultrasound. Given that women with POD obliteration at laparoscopy have an increased risk of bowel endometriosis and requirement for bowel surgery, the TVS 'sliding sign' test should be considered in the pre-operative imaging work-up for all women with suspected endometriosis, to allow for appropriate surgical planning. We believe the TVS 'sliding sign' technique may be easily learned by sonologists/sonographers who are familiar with performing gynaecological ultrasound, and that further studies are required to confirm the diagnostic accuracy of this new ultrasound technique amongst sonologists/sonographers with various levels of experience.

**STUDY FUNDING/COMPETING INTEREST(S):** This study received no specific grant from any funding agency in the public, commercial or not-for-profit sectors and the authors declare no competing interests.

**Key words:** endometriosis / transvaginal ultrasound / 'sliding sign' / pouch of Douglas obliteration

# Introduction

Transvaginal ultrasound (TVS) has been shown to be accurate in the pre-operative diagnosis of both ovarian (Mais *et al.*, 1993; Guerriero *et al.*, 1996) and extra-ovarian endometriosis, particularly with regard to deep infiltrating endometriosis (DIE) of the bowel and bladder (Bazot *et al.*, 2003, 2004, 2007; Abrao *et al.*, 2007; Hudelist *et al.*, 2009; Holland *et al.*, 2010). Bowel endometriosis affects between 3.8 and 37% of women with endometriosis (Remorgida *et al.*, 2007). The excision of DIE may be a long and complex operation, requiring the skills of an advanced laparoscopic surgeon (Chapron *et al.*, 2006). In the case of bowel endometriosis, the expertise of a colorectal surgeon is often required. Although there are published data confirming the effectiveness of TVS in the prediction of DIE (Abrao *et al.*, 2007; Bazot *et al.*, 2007; Piketty *et al.*, 2009; Hudelist *et al.*, 2009, 2011b; Holland *et al.*, 2010; Reid *et al.*, 2011a,b), very little exists in relation to predicting pouch of Douglas (POD) obliteration. Along with bowel endometriosis, POD obliteration also requires high-level operative skills. In a study by Khong *et al.* (2011), women with POD obliteration had a 3-fold higher risk of having bowel endometriosis and bowel surgery, compared with those without POD obliteration. Therefore, the ability to predict POD obliteration pre-operatively may aid in the triaging and planning for women undergoing endometriosis surgery.

A systematic review and meta-analysis by Hudelist *et al.* (2011b) found TVS to have a high sensitivity (91%) and specificity (98%) in the prediction of bowel endometriosis. However, the results of this review may be relevant only for women with a high risk of DIE, as most of the women included in the analysis were treated in tertiary referral centres for endometriosis. Recent studies have also used TVS to predict POD obliteration with a sensitivity of 72–83% and specificity of 97–100% (Reid *et al.*, 2012; Hudelist *et al.*, 2011a). One of the criticisms of these studies is that the technique described was not validated between different observers in terms of accuracy and reproducibility. There are no data available on the inter-observer agreement using TVS for the prediction of POD obliteration. The aim of this study was to determine the inter-/intra-observer agreement and the reliability of TVS 'sliding sign' in the prediction of POD obliteration (secondary to endometriosis) at offline analysis of videos using the dynamic real-time TVS 'sliding sign' technique.

# Materials and Methods

Pre-recorded video sets evaluating the POD in 30 women presenting with chronic pelvic pain, to the Endogynaecology tertiary referral clinic at Nepean Hospital, were assessed 'offline' by six observers. All videos were recorded by the same sonologist (S.R.) using 7.5 MHz transvaginal probe (Medison V20, Samsung Medison, Seoul, Republic of Korea). The videos included in this study were selected based on real-time representative quality/clarity, and were not obtained from sequential patients. The 'sliding sign' test (Reid *et al.*, 2012) and POD obliteration final

outcome for each of the 30 women were confirmed with 'gold standard' laparoscopy in 24 out of 30 (80%) cases. POD obliteration was present at laparoscopy for 6 out of 24 (25%) women, and 18 out of 24 (75%) women had no POD obliteration at laparoscopy. Of the remaining, six women (20%) who did not undergo 'gold standard' laparoscopy, five (5/6 = 83.3%) were found to have a clearly positive 'sliding sign' when their videos were reviewed by S.R., who is experienced in the prediction of POD obliteration using the real-time TVS 'sliding sign' technique. Similarly, in the remaining one woman (1/6 = 16.7%), the 'sliding sign' video clearly demonstrated a negative 'sliding sign' (again this video was reviewed by S.R.).

Each sonologist observer viewed the TVS 'sliding sign' video in two anatomical locations (retro-cervix and posterior uterine fundus), i.e. 60 videos in total were reviewed. Four of the observers performed daily gynaecological scanning, while two were primarily fetal medicine sonologists. Each sonologist was blinded to the outcome of POD obliteration. The first pre-recorded video of each of the 30 women included the assessment of the 'sliding sign' in the posterior uterine fundal region. In this video, the sonologist had placed one hand over the lower abdomen of the woman to elicit pressure on the uterine fundus while performing a TVS simultaneously. When assessing the offline video in this region, each of the six observers had to decide if there was free movement of the anterior recto-sigmoid across the posterior uterine fundus. If so, then the observer noted the result as a positive 'sliding sign'. In the second pre-recorded video for each of the 30 women, the sonologist had placed the TVS probe in the posterior fornix to assess the 'sliding sign' in the retro-cervical area. If, by placing gentle pressure on the posterior cervix, the anterior rectal wall glided smoothly over the retro-cervix, then each of the six observers noted the result as a positive 'sliding sign'. A negative 'sliding sign' video was noted by each of the six observers if the anterior recto-sigmoid colon or the anterior rectum was fixed to the posterior uterine fundus or retro-cervix, respectively.

Each observer was then asked to classify whether the POD was obliterated or not on the basis of evaluating these pre-recorded video sets. If the observer deemed that the 'sliding sign' was positive in both anatomical locations (i.e. the anterior rectum and the recto-sigmoid colon glided smoothly across the retro-cervix and posterior uterine fundus, respectively), then the POD was noted to be not obliterated. If the observer deemed that the 'sliding sign' was negative for either region (i.e. the anterior recto-sigmoid colon and/or the anterior rectum did not glide smoothly over posterior uterine fundal region and/or retro-cervix, respectively), then the POD was noted to be obliterated. Each sonologist was also asked to reanalyse the same videos, albeit in a different order, at least 7 days later to assess intra-observer agreement.

A sub-analysis was performed to assess the performance of the four gynaecological sonologists in terms of inter-observer correlation. These results were then compared with the inter-observer performance of all six observers.

## Statistical analysis

Diagnostic accuracy and inter-observer agreement for both the 'sliding sign' and POD obliteration diagnosis were evaluated for all six observers. A separate analysis was also performed for four of the observers who specialized in gynaecological ultrasound, to determine the multiple rater agreement for the interpretation of the 'sliding sign' and prediction of POD obliteration. The data analysis included pairwise rater agreement, which determined the variability between any two observers on the three categories of the 'sliding sign' outcome (yes, no and unsure) and POD obliteration outcome (yes, no and unsure), using Cohen's $\kappa$ statistics (corrected for agreement by chance) with 95% confidence intervals (CIs). General rules for the interpretation of Cohen's $\kappa$ coefficient were used, i.e. $<0$ = poor agreement, 0.01–0.20 = slight agreement, 0.21–0.40 = fair agreement, 0.41–0.60 = moderate agreement, 0.61–0.80 = substantial agreement and 0.81–0.99 = almost perfect agreement. Multiple rater agreement was analysed between all six observers, and then between the four gynaecological ultrasound specialists, for the three category 'sliding sign' and POD outcomes. Fleiss' $\kappa$ (with 95% CIs) was used for overall and individual outcomes in the multiple rater agreement analysis. The multiple rater agreement on sliding sign outcome for the two anatomical regions (posterior uterine fundus and retro-cervix) was also analysed. The diagnostic accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the six observers were analysed for both the 'sliding sign' and POD obliteration outcomes (the cases classified by the observers as 'unsure' for POD outcome were excluded from the diagnostic accuracy analysis). The intra-observer agreement for all six observers was also analysed for the interpretation of the 'sliding sign' and prediction of POD obliteration using Cohen's $\kappa$ coefficient with 95% CIs.

# Results

## Inter-observer analysis

The pairwise rater agreement (Cohen's $\kappa$) with 95% CIs for the interpretation of the 'sliding sign' (yes, no and unsure) between all six observers for both videos sets and both anatomical regions (fundal and retro-cervix) is shown in Table I. Agreement between all six observers for both sets of videos in both regions ranged from 0.354 to 0.927 (fair agreement to almost perfect agreement), compared with 0.630–0.927 (substantial agreement to almost perfect agreement) when the two fetal medicine observers were excluded.

The multiple rater agreement for the three categories for 'sliding sign' outcome (yes, no and unsure) for all six observers using Fleiss' $\kappa$ with 95% CIs is shown in Table II. The overall agreement between all six observers on the interpretation of the 'sliding sign' for both video sets and both regions (retro-cervix and fundal) was moderate (Fleiss' $\kappa$ 0.45, $P$-value $<0.01$). The overall agreement on the 'sliding sign' for the retro-cervical region was higher (0.54) compared with the fundal region (0.37), and these values were significant ($P < 0.01$). When the four observers specializing in gynaecological ultrasound were analysed separately (Table III), the overall agreement between observers on the interpretation of the 'sliding sign' was substantial (Fleiss' $\kappa$ 0.65, $P$-value $<0.01$). There was also stronger agreement for the interpretation of the 'sliding sign' in the retro-cervical region (0.73) compared with the posterior fundal region (0.56), and these values were also significant ($P < 0.01$).

Table IV includes the inter-observer pairwise agreement for the prediction of POD obliteration. Again, the four observers specializing in gynaecological ultrasound displayed stronger inter-observer agreement for the diagnosis of POD obliteration when compared with the fetal medicine specialists. When comparing all six observers for the multiple rater agreement on the three category prediction of POD obliteration, overall Fleiss' $\kappa$ was 0.427 (moderate agreement), compared with 0.694 (substantial agreement) for the four gynaecological ultrasound specialists (Table V). The $\kappa$ for agreement on the presence or absence of POD obliteration for all the six observers was 0.532 and 0.461, respectively. These $\kappa$-values improved when the subgroup of the four gynaecological ultrasound specialists was analysed, with the prediction of the presence or absence of POD

**Table I. Pairwise rater agreement for all six observers, which demonstrates the variability between any two observers on the three categories of the 'sliding sign' outcome (yes, no and unsure).**

| | Observer 1a | Observer 2b | Observer 3a | Observer 4b | Observer 5a | Observer 6a | True result |
|---|---|---|---|---|---|---|---|
| Both video sets (1 and 2) and both regions (fundal and retro-cervical regions) | | | | | | | |
| Observer 1 | N/A | 0.341 | 0.613 | 0.439 | 0.725 | 0.732 | 0.927 |
| Observer 2 | 0.21–0.47 | N/A | 0.183 | 0.11 | 0.216 | 0.352 | 0.354 |
| Observer 3 | 0.46–0.76 | 0.09–0.28 | N/A | 0.544 | 0.673 | 0.521 | 0.630 |
| Observer 4 | 0.31–0.57 | 0.04–0.18 | 0.4–0.69 | N/A | 0.446 | 0.373 | 0.423 |
| Observer 5 | 0.57–0.88 | 0.11–0.33 | 0.53–0.82 | 0.31–0.58 | N/A | 0.656 | 0.743 |
| Observer 6 | 0.57–0.9 | 0.22–0.48 | 0.38–0.66 | 0.25–0.49 | 0.51–0.8 | N/A | 0.802 |
| True result | 0.75–1 | 0.23–0.48 | 0.48–0.78 | 0.29–0.55 | 0.59–0.9 | 0.64–0.97 | N/A |
| Video set 1 on both regions (fundal and retro-cervical regions) | | | | | | | |
| Observer 1 | N/A | 0.41 | 0.643 | 0.496 | 0.648 | 0.703 | 0.902 |
| Observer 2 | 0.21–0.61 | N/A | 0.259 | 0.159 | 0.225 | 0.381 | 0.41 |
| Observer 3 | 0.42–0.86 | 0.11–0.41 | N/A | 0.594 | 0.704 | 0.603 | 0.722 |
| Observer 4 | 0.29–0.7 | 0.03–0.29 | 0.39–0.8 | N/A | 0.486 | 0.438 | 0.496 |
| Observer 5 | 0.42–0.88 | 0.05–0.4 | 0.49–0.91 | 0.29–0.68 | N/A | 0.689 | 0.692 |
| Observer 6 | 0.47–0.94 | 0.18–0.58 | 0.4–0.81 | 0.25–0.62 | 0.48–0.9 | N/A | 0.802 |
| True result | 0.65–1 | 0.21–0.61 | 0.5–0.94 | 0.29–0.7 | 0.46–0.92 | 0.57–1 | N/A |
| Video set 2 on both regions (fundal and retro-cervical regions) | | | | | | | |
| Observer 1 | N/A | 0.277 | 0.587 | 0.388 | 0.796 | 0.76 | 0.952 |
| Observer 2 | 0.12–0.43 | N/A | 0.124 | 0.07 | 0.208 | 0.324 | 0.299 |
| Observer 3 | 0.38–0.79 | 0.01–0.24 | N/A | 0.492 | 0.645 | 0.451 | 0.549 |
| Observer 4 | 0.21–0.56 | −0.02 to 0.16 | 0.3–0.69 | N/A | 0.408 | 0.316 | 0.36 |
| Observer 5 | 0.57–1 | 0.07–0.35 | 0.44–0.85 | 0.23–0.59 | N/A | 0.624 | 0.792 |
| Observer 6 | 0.53–0.99 | 0.16–0.49 | 0.26–0.64 | 0.16–0.47 | 0.42–0.83 | N/A | 0.802 |
| True result | 0.7–1 | 0.14–0.46 | 0.35–0.75 | 0.19–0.53 | 0.57–1 | 0.57–1 | N/A |
| Both video sets (1 and 2) for fundal region only | | | | | | | |
| Observer 1 | N/A | 0.348 | 0.539 | 0.294 | 0.661 | 0.674 | 1 |
| Observer 2 | 0.16–0.54 | N/A | 0.147 | 0.071 | 0.227 | 0.421 | 0.348 |
| Observer 3 | 0.34–0.74 | 0.03–0.26 | N/A | 0.475 | 0.641 | 0.395 | 0.539 |
| Observer 4 | 0.13–0.45 | −0.01 to 0.16 | 0.29–0.66 | N/A | 0.338 | 0.26 | 0.294 |
| Observer 5 | 0.44–0.88 | 0.08–0.37 | 0.44–0.84 | 0.17–0.51 | N/A | 0.522 | 0.661 |
| Observer 6 | 0.45–0.9 | 0.23–0.61 | 0.22–0.57 | 0.12–0.4 | 0.32–0.72 | N/A | 0.674 |
| True result | 0.75–1 | 0.16–0.54 | 0.34–0.74 | 0.13–0.45 | 0.44–0.88 | 0.45–0.9 | N/A |
| Both video sets (1 and 2) for retro-cervical region only | | | | | | | |
| Observer 1 | N/A | 0.333 | 0.695 | 0.624 | 0.789 | 0.778 | 0.864 |
| Observer 2 | 0.16–0.5 | N/A | 0.226 | 0.167 | 0.207 | 0.301 | 0.358 |
| Observer 3 | 0.47–0.92 | 0.08–0.37 | N/A | 0.617 | 0.707 | 0.659 | 0.73 |
| Observer 4 | 0.41–0.83 | 0.04–0.29 | 0.41–0.83 | N/A | 0.570 | 0.526 | 0.587 |
| Observer 5 | 0.56–1 | 0.04–0.37 | 0.49–0.92 | 0.37–0.77 | N/A | 0.788 | 0.828 |
| Observer 6 | 0.54–1 | 0.13–0.47 | 0.45–0.87 | 0.33–0.72 | 0.57–1 | N/A | 0.909 |
| True result | 0.61–1 | 0.18–0.53 | 0.5–0.96 | 0.38–0.79 | 0.6–1 | 0.67–1 | N/A |

The upper right entry from the diagonal is the Cohen's $\kappa$ corrected for agreement by chance. The lower left entry from the diagonal is the 95% CIs for the corresponding Cohen's $\kappa$. General rules for interpretation of $\kappa$. <0 poor agreement, 0.01–0.20 slight agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement and 0.81–0.99 almost perfect agreement. N/A, not applicable. Video set 1, the first analysis; Video set 2, the second analysis, with the re-ordering of the cases presented to the observers.

[a]Gynaecological sonologist/sonographer.

[b]Fetal medicine sonologist.

**Table II. Multiple rater agreement for all six observers on the three category (yes, no and unsure) interpretation of the 'sliding sign'.**

| Data set | Overall | Yes | No | Unsure |
|---|---|---|---|---|
| Both video sets (1 and 2) and both regions (fundal and retro-cervical regions) | 0.45 (0.41–0.49) | 0.49 (0.44–0.53) | 0.52 (0.48–0.57) | 0.03* (−0.02 to 0.07) |
| Video set 1 and both regions (fundal and retro-cervical regions) | 0.50 (0.44–0.55) | 0.52 (0.46–0.59) | 0.56 (0.49–0.62) | 0.06* (−0.01 to 0.12) |
| Video set 2 and both regions (fundal and retro-cervical regions) | 0.41 (0.41–0.49) | 0.45 (0.44–0.53) | 0.49 (0.48–0.57) | 0.01* (−0.02 to 0.07) |
| Both video sets (1 and 2) on fundal region | 0.37 (0.31–0.43) | 0.38 (0.31–0.44) | 0.45 (0.38–0.51) | 0.04* (−0.02 to 0.11) |
| Both video sets (1 and 2) on retro-cervical region | 0.54 (0.49–0.60) | 0.60 (0.53–0.66) | 0.60 (0.54–0.67) | 0.01* (−0.05 to 0.08) |

Fleiss' $\kappa$ was used for overall and individual outcomes. Within brackets are the 95% CIs for the $\kappa$-values. Video set 1, the first analysis; Video set 2, the second analysis, with the re-ordering of the cases presented to the observers.
*Not significantly different from 0 ($P$-value $>0.05$). The other $\kappa$-values without '*' marks are all significantly different from 0 with $P$-value $<0.01$.

**Table III. Multiple rater agreement for the four observers specializing in gynaecological ultrasound on the three category (yes, no and unsure) interpretation of the 'sliding sign'.**

| Data set | Overall | Yes | No | Unsure |
|---|---|---|---|---|
| Both video sets (1 and 2) and both regions (fundal and retro-cervical regions) | 0.65 (0.58–0.71) | 0.68 (0.60–0.75) | 0.75 (0.68–0.83) | 0.03* (−0.05 to 0.10) |
| Video set 1 and both regions (fundal and retro-cervical regions) | 0.66 (0.57–0.75) | 0.69 (0.58–0.79) | 0.77 (0.67–0.87) | 0.04* (−0.07 to 0.14) |
| Video set 2 and both regions (fundal and retro-cervical regions) | 0.63 (0.58–0.71) | 0.66 (0.60–0.75) | 0.74 (0.68–0.83) | 0.02* (−0.05 to 0.10) |
| Both video sets (1 and 2) on fundal region | 0.56 (0.47–0.65) | 0.57 (0.47–0.67) | 0.70 (0.60–0.81) | 0.01* (−0.10 to 0.11) |
| Both video sets (1 and 2) on retro-cervical region | 0.73 (0.64–0.82) | 0.777 (0.67–0.88) | 0.80 (0.70–0.91) | 0.05* (−0.05 to 0.16) |

Fleiss' $\kappa$ was used for overall and individual outcomes. Within brackets are the 95% CIs for the $\kappa$-values. Video set 1, the first analysis; Video set 2, the second analysis, with the re-ordering of the cases presented to the observers.
*Not significantly different from 0 ($P$-value $> 0.05$). The other $\kappa$-values without '*' marks are all significantly different from 0 with $P$-value $<0.01$.

obliteration having substantial agreements of 0.798 and 0.758, respectively.

## Intra-observer analysis

The intra-observer agreement among the six observers for both video sets and both regions in the interpretation of the 'sliding sign' is shown in Table VI. The $\kappa$-values ranged from 0.60 to 0.95 (moderate agreement to almost perfect agreement) for the interpretation of the 'sliding sign' for both anatomical regions for all six observers, whereas the intra-observer agreement was 0.71–0.95 (substantial to almost perfect) when the four gynaecological ultrasound specialists were considered. When the interpretation of the 'sliding sign' was analysed for each anatomical region for all six observers, the intra-observer reproducibility was found to be similar for the fundal and retro-cervical regions (both had moderate to substantial agreements). When the two fetal medicine observers were excluded, the intra-observer agreement for the interpretation of the 'sliding sign' was slightly higher for the retro-cervical region ($\kappa$ 0.80–0.91) compared with the fundal region ($\kappa$ 0.63–1.0), and both regions displayed substantial to almost perfect intra-observer agreement.

Table VII displays the intra-observer agreement for the prediction of POD obliteration for both video sets, which ranged from moderate to almost perfect agreement ($\kappa$ 0.46–1.0) for all six observers. The intra-observer agreement for the prediction of POD obliteration ranged from substantial to almost perfect agreement ($\kappa$ 0.67–1.0) when the subgroup of gynaecological sonologists was considered for the prediction of POD obliteration.

The reliability of the 'sliding sign' technique in the prediction of POD obliteration was evaluated by comparing the diagnostic outcomes (i.e. POD obliteration versus no obliteration) assigned by all six observers with the previously determined diagnostic outcomes for POD obliteration. As noted in the Materials and Methods section, the 'gold standard' (i.e. laparoscopy findings) for the presence/absence of POD obliteration was known for 24 out of 30 cases, whereas the POD obliteration outcome for the remaining six cases was determined by an experienced clinician by the interpretation of the 'sliding sign' (S.R.). For the prediction of POD obliteration for all six observers, after excluding cases with the POD outcome classified as 'unsure' by the observers, the TVS 'sliding sign' test gave an accuracy, sensitivity, specificity, PPV and NPV of 70.2–100, 33.3–100, 57.6–100, 50.0–100 and 85.2–100%, respectively. When comparing the four gynaecological observers (after excluding cases with the POD outcome classified as 'unsure' by the observers), the results for accuracy, sensitivity, specificity, PPV and NPV were 93.1–100, 92.9–100,

**Table IV. Pairwise rater agreement for all six observers, which demonstrates the variability between any two observers on the three category prediction (yes, no and unsure) of POD obliteration.**

| | Observer 1a | Observer 2b | Observer 3a | Observer 4b | Observer 5a | Observer 6a | True result |
|---|---|---|---|---|---|---|---|
| Both video sets 1 and 2 | | | | | | | |
| Observer 1 | N/A | 0.379 | 0.768 | 0.352 | 0.792 | 0.657 | 0.911 |
| Observer 2 | 0.2–0.56 | N/A | 0.296 | 0.089 | 0.278 | 0.246 | 0.434 |
| Observer 3 | 0.53–1 | 0.13–0.46 | N/A | 0.377 | 0.697 | 0.614 | 0.761 |
| Observer 4 | 0.2–0.5 | 0–0.18 | 0.21–0.54 | N/A | 0.309 | 0.351 | 0.306 |
| Observer 5 | 0.55–1 | 0.1–0.46 | 0.47–0.92 | 0.16–0.46 | N/A | 0.659 | 0.826 |
| Observer 6 | 0.46–0.86 | 0.08–0.41 | 0.42–0.81 | 0.2–0.5 | 0.46–0.86 | N/A | 0.723 |
| True result | 0.66–1 | 0.24–0.62 | 0.53–0.99 | 0.16–0.45 | 0.59–1 | 0.52–0.92 | N/A |
| Video set 1 | | | | | | | |
| Observer 1 | N/A | 0.468 | 0.83 | 0.404 | 0.686 | 0.623 | 0.911 |
| Observer 2 | 0.17–0.77 | N/A | 0.468 | 0.149 | 0.231 | 0.245 | 0.535 |
| Observer 3 | 0.47–1 | 0.17–0.77 | N/A | 0.35 | 0.764 | 0.698 | 0.911 |
| Observer 4 | 0.18–0.63 | 0–0.3 | 0.12–0.58 | N/A | 0.364 | 0.379 | 0.353 |
| Observer 5 | 0.37–1 | −0.03 to 0.49 | 0.45–1 | 0.13–0.59 | N/A | 0.567 | 0.674 |
| Observer 6 | 0.35–0.9 | 0.01–0.48 | 0.43–0.97 | 0.16–0.6 | 0.3–0.84 | N/A | 0.686 |
| True result | 0.55–1 | 0.22–0.85 | 0.55–1 | 0.13–0.57 | 0.36–0.99 | 0.41–0.96 | N/A |
| Video set 2 | | | | | | | |
| Observer 1 | N/A | 0.295 | 0.717 | 0.305 | 0.911 | 0.692 | 0.911 |
| Observer 2 | 0.09–0.5 | N/A | 0.175 | 0.038 | 0.338 | 0.252 | 0.338 |
| Observer 3 | 0.41–1 | 0–0.35 | N/A | 0.392 | 0.639 | 0.539 | 0.639 |
| Observer 4 | 0.11–0.5 | −0.07 to 0.15 | 0.16–0.62 | N/A | 0.264 | 0.324 | 0.264 |
| Observer 5 | 0.55–1 | 0.12–0.55 | 0.34–0.94 | 0.08–0.45 | N/A | 0.76 | 1 |
| Observer 6 | 0.4–0.99 | 0.04–0.47 | 0.26–0.82 | 0.12–0.52 | 0.47–1 | N/A | 0.76 |
| True result | 0.55–1 | 0.12–0.55 | 0.34–0.94 | 0.08–0.45 | 0.64–1 | 0.47–1 | N/A |

The upper right entry from the diagonal is the Cohen's $\kappa$ corrected for agreement by chance. The lower left entry from the diagonal is the 95% CIs for the corresponding Cohen's $\kappa$. Video set 1, the first analysis; Video set 2, the second analysis, with the re-ordering of the cases presented to the observers. N/A, not applicable.
[a]Gynaecological sonologist/sonographer.
[b]Fetal medicine sonologist.

**Table V. Prediction of POD obliteration using the TVS 'sliding sign': multiple rater agreement using Fleiss' $\kappa$ for all six observers compared with the four gynaecological ultrasound specialists.**

| | All six observers | Four gynaecological ultrasound specialists |
|---|---|---|
| Overall prediction of POD obliteration | 0.427 (0.374–0.481) | 0.694 (0.605–0.783) |
| Presence of POD obliteration | 0.532 (0.467–0.598) | 0.798 (0.694–0.901) |
| Absence of POD obliteration | 0.461 (0.395–0.526) | 0.758 (0.655–0.861) |

Within brackets are the 95% CIs for the $\kappa$-values. POD, pouch of Douglas; TVS, transvaginal ultrasound.

90.9–100, 77.8–100 and 97.7–100%, respectively (Table VIII). Table VIII also includes the number and rate (%) of 'unsure' cases for the prediction of POD obliteration for each of the six observers.

## Discussion

This is the first study to validate the TVS 'sliding sign' technique for the prediction of POD obliteration. These results demonstrate the high reproducibility of this ultrasound-based method. This was especially so for the gynaecological sonologists whose inter- and intra-observer results demonstrated substantial to almost perfect agreement. A similar pattern was also noted among the gynaecological sonologists when determining POD obliteration based upon video analysis of the posterior compartment. When all six observers were grouped together, the results were still encouraging, although not as impressive.

**Table VI. Intra-observer variability on the interpretation of the 'sliding sign' for both regions and individual regions for all six observers.**

| | Observer 1a | Observer 2b | Observer 3a | Observer 4b | Observer 5a | Observer 6a |
|---|---|---|---|---|---|---|
| Both regions (fundal and retro-cervical regions) | | | | | | |
| Agree on yes | 46 | 55 | 33 | 23 | 40 | 48 |
| Agree on no | 13 | 2 | 16 | 21 | 11 | 9 |
| Agree on unsure | 0 | 0 | 2 | 2 | 2 | 1 |
| Disagree | 1 | 3 | 9 | 14 | 7 | 2 |
| $\kappa$ | 0.95 | 0.61 | 0.71 | 0.60 | 0.73 | 0.90 |
| 95% CI | 0.70–1 | 0.41–0.80 | 0.51–0.92 | 0.40–0.80 | 0.52–0.94 | 0.68–1 |
| Fundal region only | | | | | | |
| Agree on yes | 24 | 28 | 15 | 9 | 19 | 25 |
| Agree on no | 6 | 1 | 8 | 13 | 5 | 4 |
| Agree on unsure | 0 | 0 | 1 | 1 | 1 | 1 |
| Disagree | 0 | 1 | 6 | 7 | 5 | 0 |
| $\kappa$ | 1.00 | 0.65 | 0.64 | 0.60 | 0.63 | 1.00 |
| 95% CI | 0.64–1 | 0.32–0.92 | 0.36–0.92 | 0.32–0.88 | 0.34–0.92 | 0.69–1 |
| Retro-cervical region only | | | | | | |
| Agree on yes | 22 | 27 | 18 | 14 | 21 | 23 |
| Agree on no | 7 | 1 | 8 | 8 | 6 | 5 |
| Agree on unsure | 0 | 0 | 1 | 1 | 1 | 0 |
| Disagree | 1 | 2 | 3 | 7 | 2 | 2 |
| $\kappa$ | 0.91 | 0.58 | 0.80 | 0.58 | 0.84 | 0.82 |
| 95% CI | 0.55–1 | 0.34–0.82 | 0.49–1 | 0.30–0.86 | 0.54–1 | 0.55–1 |

The first four rows for each category report the number of agreements on different outcome categories (between the first and second set of diagnosis). CI, confidence interval.
[a]Gynaecological sonologist/sonographer.
[b]Fetal medicine sonologist.

**Table VII. Intra-observer variability on prediction of POD obliteration, displaying the number of agreements on different diagnosis categories (yes, no and unsure) for the six different observers (between the first and second set of diagnosis).**

| | Observer 1a | Observer 2b | Observer 3a | Observer 4b | Observer 5a | Observer 6a |
|---|---|---|---|---|---|---|
| Agree on yes | 8 | 1 | 8 | 12 | 6 | 4 |
| Agree on no | 22 | 26 | 18 | 7 | 20 | 21 |
| Agree on unsure | 0 | 0 | 0 | 4 | 0 | 2 |
| Disagree | 0 | 3 | 4 | 7 | 4 | 3 |
| $\kappa$ | 1 | 0.46 | 0.72 | 0.64 | 0.67 | 0.78 |
| 95% CI | 0.64–1 | 0.22–0.71 | 0.41–1 | 0.38–0.89 | 0.36–0.99 | 0.51–1 |

All $\kappa$-values in the table are significantly different from 0 with a $P$-value of $<0.01$. CI, confidence interval; POD, pouch of Douglas.
[a]Gynaecological sonologist/sonographer.
[b]Fetal medicine sonologist.

Overall inter-observer multiple rater agreement for the prediction of the 'sliding sign' outcome was moderate for the six observers and substantial for the subgroup analysis (four specialists in gynaecological ultrasound), which indicates that experience with gynaecological ultrasound may influence the observer's ability to interpret the 'sliding sign' as positive or negative. Both groups of observers (gynaecological sonologists and gynaecological sonologists/fetal medicine specialists) found the interpretation of the posterior uterine fundal region more difficult than the retro-cervical region, as indicated by the discrepancy in inter-observer agreement for these regions.

The ability to assess the posterior uterine fundal region using the 'sliding sign' can be more challenging than the retro-cervical region. For some women, it can be difficult to elicit a positive 'sliding sign' by balloting the uterus through the abdominal wall (between the left hand and the tip of the transvaginal probe). This, in turn, may have resulted in these 'sliding sign' videos of the posterior fundal region

**Table VIII. Performance of all six individual observers on prediction of POD obliteration after viewing the 'sliding sign' technique, for both video sets (after exclusion of unsure cases).**

| | Observer 1a | Observer 2b | Observer 3a | Observer 4b | Observer 5a | Observer 6a |
|---|---|---|---|---|---|---|
| Accuracy (%) | 96.67 | 86.21 | 93.10 | 70.21 | 96.55 | 100.00 |
| Sensitivity (%) | 100.00 | 33.33 | 100.00 | 100.00 | 92.86 | 100.00 |
| Specificity (%) | 95.65 | 100.00 | 90.91 | 57.58 | 97.73 | 100.00 |
| PPV (%) | 87.50 | 100.00 | 77.78 | 50.00 | 92.86 | 100.00 |
| NPV (%) | 100.00 | 85.19 | 100.00 | 100.00 | 97.73 | 100.00 |
| Number of unsure cases | 0 | 2 | 2 | 13 | 2 | 7 |
| Rate of unsure cases (%) | 0.00 | 3.33 | 3.33 | 21.67 | 3.33 | 11.67 |

PPV, positive predictive value; NPV, negative predictive value; POD, pouch of Douglas.
[a]Gynaecological sonologist/sonographer.
[b]Fetal medicine sonologist.

being more difficult to interpret. When the 'sliding sign' is used to assess the retro-cervical region, the TV probe is in direct contact with the uterine cervix, and movement of the uterus is more easily elicited. Therefore, the retro-cervical region may be more easily assessed for the gliding of the anterior rectum over the retro-cervix and posterior vaginal wall. This difference in ability to mobilize the uterus for the two anatomical regions may explain the decrease in inter-observer agreement for the posterior fundal region compared with the retro-cervical region. The observers who specialized in gynaecological ultrasound had moderate to almost perfect inter-observer agreement for interpretation of the 'sliding sign' in either anatomical region, whereas the agreement for the fetal medicine sonologists ranged from only fair to moderate, indicating that experience with gynaecological ultrasound is likely to influence the ability to interpret whether the sliding sign is positive or negative.

Inter-observer agreement for the diagnosis of POD obliteration was based on the previously assigned 'sliding sign' outcomes for the two anatomical regions (posterior uterine fundus and retro-cervix). If the 'sliding sign' was negative in either anatomical region, the POD was recorded as obliterated. The four observers who perform gynaecological ultrasound regularly had a stronger overall inter-observer agreement for the prediction of POD obliteration ($\kappa$ 0.69) compared with all six observers ($\kappa$ 0.43). Again, this difference between the main group and subgroup may be explained by the fact that the observers who specialize in transvaginal gynaecological ultrasound have a greater familiarity with interpreting gynaecological ultrasound findings. In addition, the gynaecological ultrasound specialists were able to correctly classify both the presence and absence of POD obliteration with substantial agreement ($\kappa$ 0.798 and 0.758, respectively). The ability to predict for the presence or absence of POD obliteration was higher for the four observers who specialize in gynaecological ultrasound compared with the two fetal medicine specialists, with certainty of diagnosis ranging from 70.0 to 93.0% versus 45.0 to 88.3%, respectively. This finding is also likely to be a reflection of the gynaecological sonologists' increased experience and confidence with interpreting gynaecological ultrasound findings.

The intra-observer agreement for the six observers ranged from moderate to nearly perfect agreement on both video sets for the interpretation of the 'sliding sign'. For the gynaecological sonologists, $\kappa$-values were consistent with substantial to nearly perfect agreement. When interpreting the 'sliding sign' for the two anatomical regions, the intra-observer agreement was slightly higher for the fundal versus retro-cervical region for the two fetal medicine specialists. The opposite was true for the four specialists in gynaecological ultrasound, where the retro-cervical region had a slightly higher intra-observer agreement than the fundal region.

Pre-operative TVS has been used to predict POD obliteration with a reported sensitivity of 72–83% and specificity of 97–100% (Holland *et al.*, 2010; Hudelist *et al.*, 2011a; Reid *et al.*, 2012). The use of the dynamic, real-time TVS 'sliding sign' technique to systematically assess the posterior compartment for POD obliteration appears to allow for the accurate diagnosis of POD obliteration. In the current study, the four observers specializing in gynaecological ultrasound were found to have an accuracy, sensitivity, specificity, PPV and NPV ranging from 93.1 to 100, 92.9 to 100, 90.9 to 100, 77.8 to 100 and 97.7 to 100%, respectively, when the TVS 'sliding sign' videos were used to interpret whether the POD was obliterated.

The diagnostic accuracy findings should be interpreted with the caveat that the cases classified as 'unsure' for the prediction of POD obliteration were excluded from the analysis, which means that the performance of each observer is only relevant for the number of cases that the observer classified as 'yes' or 'no' for POD obliteration. Although the 'unsure' cases for the prediction of POD obliteration were excluded in the diagnostic accuracy analysis, the rate of 'unsure' cases was not excessively high, ranging from 0 to 11.7% for the gynaecological ultrasound specialists (compared with 3.33–21.7% for the two fetal medicine specialists). Our ability to draw conclusions regarding the diagnostic accuracy of the TVS 'sliding sign' for POD obliteration may also be limited by the fact that the 'gold standard' laparoscopic findings were only available for 24 out of 30 (80%) cases. The remaining six cases who did not undergo laparoscopy to confirm POD outcome had POD outcomes predicted by a sonologist (S.R.) who is experienced in the prediction of POD obliteration using the TVS 'sliding sign' technique, and who also accurately predicted POD outcome pre-operatively in 100% of the 24 out of 30 women who underwent laparoscopy in this study.

This reproducibility study could have been strengthened if more observers were included, particularly sonologists/sonographers

specializing in gynaecological ultrasound. In order for a reproducibility study to accurately assess inter-observer agreement, there should be sufficient experience among the observers to make an accurate diagnosis using the diagnostic tool being studied. For this study, two of the observers were not specialized in gynaecological ultrasound. As a result, the intra- and inter-observer differences were more variable among the six observers compared with the four sonologists who specialized in gynaecological ultrasound. In addition, measures of overall agreement (i.e. Cohen's $\kappa$) assume that the raters are interchangeable. We, therefore, performed a sub-analysis for the four observers that specialized in gynaecological ultrasound. This subgroup consistently performed better than the non-gynaecological sonologists, in both intra- and inter-observer interpretation of the 'sliding sign' and prediction of POD obliteration, as well as in diagnostic accuracy.

## Conclusion

We have validated the dynamic real-time TVS 'sliding sign' technique for the prediction of POD obliteration. The high inter- and intra-observer correlation between gynaecological sonologists in evaluating the 'sliding sign' for the prediction of POD obliteration ranged from substantial to almost perfect agreement. Observers who specialized in gynaecological ultrasound had markedly better results than their fetal medicine colleagues, not only when evaluating the 'sliding sign', but also when predicting POD obliteration. The 'sliding sign' technique for the prediction of POD obliteration appears to have very acceptable diagnostic accuracy and inter-/intra- observer agreement, and should be considered in the pre-operative imaging work-up for all women with suspected endometriosis. The ability to accurately predict POD obliteration pre-operatively with TVS has major implications for the surgical planning and counselling of women with chronic pelvic pain/suspected endometriosis prior to their surgery. These women may now have a reliable pre-operative imaging test (i.e. TVS assessment of the 'sliding sign' to predict POD obliteration), which may not only tell the clinician whether the POD is obliterated, but also allow for the identification of women who may be at increased risk of requiring bowel surgery for bowel endometriosis.

## Authors' roles

S.R.: substantial contributions to conception, design and interpretation of data, drafting of the article and article revision and final approval of the version to be published. C.L.: substantial contributions to data analysis and interpretation of data, article revision and final approval of the version to be published. I.C.: substantial contributions to acquisition of data, article revision and final approval of the version to be published. B.M.: substantial contributions to acquisition of data, article revision and final approval of the version to be published. R.M.: substantial contributions to acquisition of data, article revision and final approval of the version to be published. J.L.: substantial contributions to acquisition of data, article revision and final approval of the version to be published. R.B.: substantial contributions to acquisition of data, article revision and final approval of the version to be published. G.C.: substantial contributions to conception and design, acquisition of data, interpretation of data, article revision and final approval of the version to be published.

## Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

## Conflict of interest

The authors declare that there is no conflict of interest.

## References

Abrao MS, Goncalves MO, Dias JA Jr, Podgaec S, Chamie LP, Blasbalg R. Comparison between clinical examination, transvaginal sonography and magnetic resonance imaging for the diagnosis of deep endometriosis. *Hum Reprod* 2007;**22**:3092–3097.

Bazot M, Detchev R, Cortez A, Amouyal P, Uzan S, Darai E. Transvaginal sonography and rectal endoscopic sonography for the assessment of pelvic endometriosis: a preliminary comparison. *Hum Reprod* 2003; **18**:1686–1692.

Bazot M, Thomassin I, Hourani R, Cortez A, Darai E. Diagnostic accuracy of transvaginal sonography for deep pelvic endometriosis. *Ultrasound Obstet Gynecol* 2004;**24**:180–185.

Bazot M, Malzy P, Cortez A, Roseau G, Amouyal P, Darai E. Accuracy of transvaginal sonography and rectal endoscopic sonography in the diagnosis of deep infiltrating endometriosis. *Ultrasound Obstet Gynecol* 2007;**30**:994–1001.

Chapron C, Chopin N, Borghese B, Foulot H, Dousset B, Vacher-Lavenu MC, Vieira M, Hasan W, Bricou A. Deeply infiltrating endometriosis: pathogenetic implications of the anatomical distribution. *Hum Reprod* 2006;**21**:1839–1845.

Guerriero S, Ajossa S, Paoletti AM, Mais V, Angiolucci M, Melis GB. Tumor markers and transvaginal ultrasonography in the diagnosis of endometrioma. *Obstet Gynecol* 1996;**88**:403–407.

Holland TK, Yazbek J, Cutner A, Saridogan E, Hoo WL, Jurkovic D. Value of transvaginal ultrasound in assessing severity of pelvic endometriosis. *Ultrasound Obstet Gynecol* 2010;**36**:241–248.

Hudelist G, Ballard K, English J, Wright J, Banerjee S, Mastoroudes H, Thomas A, Singer CF, Keckstein J. Transvaginal sonography vs. clinical examination in the preoperative diagnosis of deep infiltrating endometriosis. *Ultrasound Obstet Gynecol* 2011a;**37**:480–487.

Hudelist G, English J, Thomas AE, Tinelli A, Singer CF, Keckstein J. Diagnostic accuracy of transvaginal ultrasound for non-invasive diagnosis of bowel endometriosis: systematic review and meta-analysis. *Ultrasound Obstet Gynecol* 2011b;**37**:257–263.

Hudelist G, Oberwinkler KH, Singer CF, Tuttlies F, Rauter G, Ritter O, Keckstein J. Combination of transvaginal sonography and clinical examination for preoperative diagnosis of pelvic endometriosis. *Hum Reprod* 2009;**24**:1018–1024.

Khong SY, Bignardi T, Luscombe G, Lam A. Is pouch of Douglas obliteration a marker of bowel endometriosis? *J Minim Invasive Gynecol* 2011;**18**:333–337.

Mais V, Guerriero S, Ajossa S, Angiolucci M, Paoletti AM, Melis GB. The efficiency of transvaginal ultrasonography in the diagnosis of endometrioma. *Fertil Steril* 1993;**60**:776–780.

Piketty M, Chopin N, Dousset B, Millischer-Bellaische AE, Roseau G, Leconte M, Borghese B, Chapron C. Preoperative work-up for patients with deeply infiltrating endometriosis: transvaginal ultrasonography must definitely be the first-line imaging examination. *Hum Reprod* 2009; **24**:602–607.

Reid S, Winder S, Condous G. Sonovaginography: redefining the concept of a 'normal pelvis' on transvaginal ultrasound pre-laparoscopic intervention for suspected endometriosis. *Aust J Ultrasound Med* 2011a;**14**:21–24.

Reid S, Winder S, Reid G, Condous G. Office sonovaginography: redefining the concept of a normal pelvis on transvaginal ultrasound in women with suspected endometriosis. In: 21st World Congress on Ultrasound in Obstetrics and Gynecology, **Vol. 21**, 2011b. Los Angeles, CA: Wiley-Blackwell.

Reid S, Lu C, Casikar I, Reid G, Abbott J, Cario G, Chou D, Kowalski D, Cooper M, Condous G. Can we predict pouch of Douglas obliteration in women with suspected endometriosis using a new real-time dynamic transvaginal ultrasound technique: the 'sliding sign'. *Ultrasound Obstet Gynecol* 2012. doi: 10.1002/uog.12305. [Epub ahead of print].

Remorgida V, Ferrero S, Fulcheri E, Ragni N, Martin DC. Bowel endometriosis: presentation, diagnosis, and treatment. *Obstet Gynecol Surv* 2007;**62**:461–470.